April 2019

Mapping the domain structure of Ribonuclease E in Mycobacterium smegmatis

Oscar Lee
Worcester Polytechnic Institute

WORCESTER POLYTECHNIC INSTITUTE

Mapping the domain structure of Ribonuclease E in *Mycobacterium smegmatis*

Major Qualifying Project

Author:
Oscar Lee

Date Submitted:
Apr/25/2019

Report Submitted to
Professor Scarlet Shell, advisor
Professor José M. Argüello, co-advisor
Department of Chemistry and Biochemistry

This is a Major Qualifying Project submitted to the Department of Chemistry and Biochemistry, Worcester Polytechnic Institute, 100 Institute Road, Worcester, MA 01609 for partial fulfillment of the requirements for the Degree of Bachelor of Science in Biochemistry.

ABSTRACT

Mycobacterium tuberculosis is a bacterium that causes tuberculosis. Regulation of gene expression and mRNA degradation is important for the ability of the bacteria to survive the stressors of infection. A multiprotein complex called the degradosome is important for mRNA degradation in better-studied bacteria such as *E. coli*, where Ribonuclease E (RNase E) cleaves mRNA into fragments while using a scaffold domain to interact with other proteins involved in mRNA degradation. We used *M. smegmatis*, a non-pathogen relative of *M. tuberculosis*, as a model to define the putative scaffold domains of mycobacterial RNase E. Our results suggest that RNase E has two scaffold domains flanking a putative catalytic domain 493 amino acids in length.

ACKNOWLEDGMENTS

I would like to thank Dr. Scarlet Shell and Ying Zhou for guiding this project and showing tremendous support throughout its completion. I would also like to thank Diego Vargas Blanco, Dr. Carla Martini, and Huaming Sun for all the help and giving me the opportunity to work with them in the past two academic years. Thanks to Dr. Lou Roberts for helpful feedback on this report. This project was funded by an NSF grant: NSF CAREER #1652756 (to SSS).

INTRODUCTION

Tuberculosis (TB) has become one of the top ten leading causes of death, ranked even higher than human immunodeficiency virus (HIV). In 2017, the World Health Organization estimated that 10 million people developed TB and 1.6 million died from TB (WHO, 2018). Though antibiotics are effective for treating TB, long-term use of antibiotics has caused an evolution of the pathogen, through which many strains of TB developed antibiotic resistance. For the treatment of TB, people with active TB that have not yet acquired any resistance must be treated with intensive multidrug treatment for at least 6 months to eliminate the infection and avoid development of antibiotic resistance in any remaining bacteria (CDC, 2018). This situation points to an urgent need for new treatments that can cure TB more efficiently without creating further drug resistance.

Mycobacterium tuberculosis is a bacterium that causes tuberculosis. Regulation of gene expression and mRNA degradation is important for the ability of the bacteria to survive the stressors of infection. A complex known as the degradosome plays an important role in bacterial mRNA degradation. The components of the degradosome vary among species, although RNase E, PNPase, and RhlB have important roles in the process of mRNA degradation across many species of bacteria (Kovacs et al, 2013). However, there are exceptions; in *B. subtilis,* for example, RNase E is absent and RNase Y associates with other proteins to form degradosomes (Lehnik-Habrink et al, 2010).

RNase E is an abundant and essential endonuclease for many bacteria and was found to be the central component in RNA degradosome in *E. coli* (Mackie, 2013). RNase E has a scaffold domain that mediates interactions with other RNA degradation proteins to form the degradosome (Hammarlöf et al, 2015; Vanzo et al, 1998). In *E. coli*, RNase E cleaves mRNA into fragments by loading onto a single-stranded region of the RNA and scanning to find A/U rich regions, where it preferentially cleaves (Anderson et al, 2009; Kime et al, 2010). The mRNA fragments will then be cleaved into nucleotides by PNPase while the helicase RhlB unwinds any secondary structure (Mohanty et al, 2000; Coburn et al, 1999).

Degradosomes may contain additional components. In *E. coli,* for example, RNase E-scaffolded

degradosomes also contain enolase (Miczak, 1996). In addition, *E. coli* RNase E has a domain called segment A that is responsible for membrane binding, and two arginine-rich domains that were suspected to be responsible for RNA binding (Murashko et al, 2012; Khemici et al, 2008). In *C. crescentus*, the RNase E scaffold domain contains binding sites for aconitase, RNase D, and PNPase (Voss et al, 2014). However, in *C. crescentus* RhlB was found to bind to the catalytic domain of RNase E (Voss et al, 2014). In *M. tuberculosis*, the major degradosome components were recently reported to be PNPase, RhlE, RNase E and RNase J (Płociński et al, 2019). In *M. bovis* BCG (Bacillus Calmette–Guérin), purified RNase E was reported to associate with GroEL, Ppnk, and an acetyltransferase (Kovacs et al, 2005). In *M. smegmatis*, relA, MecA, GroEL, EF-Tu, DIM1, and rplB were identified as associating with RNase E (Csanadi et al, 2009).

The main purpose of this study was to use *M. smegmatis*, which is a non-pathogenic relative of *M. tuberculosis*, to map the RNase E domain structure to identify scaffold domains and construct a set of truncated RNase E strains based on the mapping results. By understanding the domain structure of *M. smegmatis* RNase E and having this set of truncated RNase E strains, it will be possible to identify the protein candidates that might interact with RNase E, and map out the region of the non-catalytic domain RNase E that each candidate protein interacts with in *M. smegmatis*.

MATERIALS & METHODS

Q5 (High Fidelity) Polymerase Chain Reaction (PCR)

All reactions followed the instruction from New England Biolabs, and the optional 5X Q5 High GC Enhancer was added. All PCR using Q5 started with 2 minutes of initial denaturation at 98 °C. After initial denaturation, the reaction was repeated for 30 cycles of denaturation for 10 seconds at 98 °C, annealing for 30 seconds at the appropriate annealing temperature depending on the primer set that was used (see Table 1), and elongation at 72 °C for 30 seconds/kb of the length of the sequence to be amplified. The reaction was then incubated for 5 minutes at 72 °C as final elongation step and stored at 4 °C until used. All PCR products were mixed with 6X DNA loading dye and loaded into 1% TAE agarose gel from Apex to run gel electrophoresis at 140V. The gel with target bands were purified using gel extraction kit and instruction from New England Biolabs.

Table 1: Primer sets with description and annealing temperatures that were used for PCR.

Description	**Primer Set ($\frac{Foward\ Primer}{Reverse\ Primer}$)**		**Annealing Temperature (°C)**
Amplifying 6X-His-3X-Flag tagged *rne* from pSS267 to make pSS337.	SSS1237	CTACAAACTCTTCCTGTCGTCATA TGAGATCGCCCTGTGGTTCCC	72
	SSS1238	TCAGTTAACTACGTCGACATCGAT ACTAGTCGTGGCTGGGCGGC	
Amplifying the backbone from pSS221 to make pSS337.	SSS1399	TATCGATGTCGACGTAGTTAAC	61
	SSS1488	ATATGACGACAGGAAGAGTT	
Amplifying tagged *rne* with a deletion of the first 145 amino acids from pSS337 for self-assembly to make pSS343.	SSS1562	ACGAGGACTGCGCCGCGGAA	72
	SSS1563	GACTGCGCCGCGGAAGCCGCCGC CGCCCTGGAAGT	

Table 1 continued

Description	Primer Set ($\frac{Foward\ Primer}{Reverse\ Primer}$)		Annealing Temperature (°C)
Amplifying tagged *rne* with a deletion of the first 308 amino acids from pSS337 for self-assembly to make pSS370.	SSS1612	GCCGTCTCGCCTGACGACCC	72
	SSS1613	GTCAGGCGAGACGGCGCCGCCG CCGCCCTGGAAGT	
Amplifying tagged *rne* with a deletion of the first 329 amino acids from pSS337 for self-assembly to make pSS371.	SSS1610	GACAAGAGCGACGACTCCGA	70
	SSS1611	GTCGTCGCTCTTGTCGCCGCCGC CGCCCTGGAAGT	
Amplifying tagged *rne* with deletions of the first 145 and last 213 amino acids for assembly with pSS221 backbone to make pSS348.	SSS1562	ACGAGGACTGCGCCGCGGAA	72
	SSS1567	CTACGAGTCGGACTTGCGCC	
Amplifying tagged *rne* with deletions of the first 329 and last 213 amino acids for assembly with pSS221 backbone to make pSS380.	SSS1610B	GACAAGAGCGACGACTCCGA	70
	SSS1567	CTACGAGTCGGACTTGCGCC	
Amplifying tagged *rne* with deletions of the first 308 and last 213 amino acids for assembly with pSS221 backbone to make pSS379.	SSS1612B	GCCGTCTCGCCTGACGACCC	72
	SSS1567	CTACGAGTCGGACTTGCGCC	
Amplifying the pSS221 backbone for insertion tagged *rne* with deletions of the first 308 and last 213 amino acids to make pSS379.	SSS1566	AAGTCCGACTCGTAGTATCGATGT CGACGTAGTTA	72
	SSS1613B	GTCAGGCGAGACGGCGCCGCCG CCGCCCTGGAAGT	

When standard PCR conditions failed to yield the desired products, touchdown PCR was used. All touchdown PCRs using Q5 began with 1 minute of initial denaturation at 98 °C. The reaction cycle of

denaturation for 10 seconds at 98 °C, annealing for 30 seconds at 10 °C above the calculated appropriate annealing temperature (see Table 2) with 1 °C lowered in each cycle, and elongation at 72 °C for 30 seconds/kb of the length of sequence amplified, was repeated for 15 cycles. The reaction was then repeated for 30 cycles of denaturation for 10 seconds at 98 °C, annealing for 30 seconds at 5 °C below the calculated appropriate annealing temperature (see Table 2), and elongation at 72 °C for 30 seconds/kb of the length of sequence amplified. The reaction was then incubated for 5 minutes at 72 °C as final elongation step and stored at 4 °C until used. Appropriately sized bands were purified by following the procedure for using the gel extraction kit from New England Biolabs.

Table 2: Primer sets with descriptions and annealing temperatures that were used for touchdown PCR.

Description	**Primer Set ($\frac{Foward\ Primer}{Reverse\ Primer}$)**		**Lowest annealing temperature (°C)**
Amplifying tagged *rne* with a deletion of the last 213 amino acids from pSS337 for self-assembly to make pSS351.	SSS1566	AAGTCCGACTCGTAGTATCGATGT CGACGTAGTTA	60
	SSS1567	CTACGAGTCGGACTTGCGCC	
Amplifying pSS221 backbone to assemble with tagged *rne* with deletions of the first 145 and last 213 amino acids to make pSS348.	SSS1566	AAGTCCGACTCGTAGTATCGATGT CGACGTAGTTA	60
	SSS1563	GACTGCGCCGCGGAAGCCGCCGC CGCCCTGGAAGT	
Amplifying pSS221 backbone to assemble with tagged *rne* with deletions of the first 329 and last 213 amino acids make pSS380.	SSS1566	AAGTCCGACTCGTAGTATCGATGT CGACGTAGTTA	72
	SSS1611B	GTCAGGCGAGACGGCGCCGCCGC CGCCCTGGAAGT	

Plasmid constructions through Gibson Assembly

All reactions used Hi-Fi DNA Assembly Master Mix and followed the procedure from New England Biolabs, except that reaction volumes were scaled down to a final volume of 5 µL. The vector mass for each reaction was 50 ng and the mass ratio of insert and vector was 1:3. All reactions were incubated at 50 °C for 1 hour and stored at -20 °C until used.

E. coli transformations

DH5-Alpha *E. coli* cells were thawed on ice and split into tubes with minimum of 20 µL of cells. 2 µL of Gibson Assembly product was added to the cells and incubated on ice for 30 minutes. The cells were then heat shocked for 30 seconds at 42 °C and incubated on ice for 5 minutes. The cells were then recovered in 200 µL of SOC medium from New England BioLabs and incubated for 1 hour at 37 °C at 200 RPM, then plated on LB agar containing appropriate antibiotics. *E. coli* colonies from plates were grown in LB broth, shaking at 200 RPM at 37 °C for 12 to 15 hours. The transformed plasmids (see Table 3) were harvested using ZR Plasmid Miniprep-Classic kit from Zymo Research using the protocol provided, and stored at -20 °C for further use.

Table 3: All plasmids that were used in this report and their *E. coli* strains

Plasmid	***E. coli* strain**	**Description**	**Antibiotics**
pSS221	SS-E_0115	Plasmid include 6X-His-3X-Flag tagged rne	Nourseothricin
pSS267	SS-E_0158	Plasmid contains attP core sequence and nourseothricin resistance marker	Nourseothricin
pSS337	SS-E_0240	Tagged RNase E with pSS221 backbone.	Nourseothricin
pSS343	SS-E_0254	First 145 amino acids deleted tagged RNase E with pSS221 backbone.	Nourseothricin
pSS348	SS-E_0258	First 145 and last 213 amino acids deleted tagged RNase E with pSS221 backbone.	Nourseothricin
pSS351	SS-E_0261	Last 213 amino acids deleted tagged RNase E with pSS221 backbone.	Nourseothricin

Table 3 continued

Plasmid	*E. coli* strain	Description	Antibiotics
pSS370	SS-E_0287	First 308 deleted tagged RNase E with pSS221 backbone.	Nourseothricin
pSS371	SS-E_0288	First 329 deleted tagged RNase E with pSS221 backbone.	Nourseothricin
pSS379	SS-E-0296	First 308 and last 213 amino acids deleted tagged RNase E with pSS221 backbone.	Nourseothricin
pSS380	SS-E-0297	First 329 and last 213 amino acids deleted tagged RNase E with pSS221 backbone.	Nourseothricin

Tagged RNase E integrated in L5 site of M. smegmatis chromosome

The *M. smegmatis* strain used in this project was SS-M_0461. Unlike the wild type of *M. smegmatis*, the original copy of RNase E located at native site was removed and the tagged RNase E was integrated into its L5 site of chromosome. To integrate the tagged RNase E into the L5 site of the chromosome, a plasmid with a backbone that contain *attP* core sequence and which includes a start codon, native promoter and 5'UTR, and 6xHis-FLAG tagged RNase E was constructed and transformed into wild type *M. smegmatis*. The *attP* core sequence in the plasmid cooperates with the *attB* sequence in L5 site of chromosome and results in a cross-over which integrates 6xHis-FLAG tagged RNase E in the L5 site.

For the removal of the original copy of RNase E in the native site, a two-step cross-over was performed. A plasmid which includes upstream sequence of RNase E, the last 150 nucleotides of the native copy RNase E, and downstream sequence of RNase E was constructed and transformed into *M. smeg*. There are two possiblities for where the first cross-over occurs as the plasmid was transformed- one upstream and the other downstream of the full-length RNaseE sequence.

After the first cross-over occurs, two copies of upstream and downstream sequences are present in the chromosome, triggering the second cross-over event. There will also be two possible sites for the

second cross-over, one upstream and the other downstream of the RNase E sequence. Two scenarios exist for the second cross-over; one is where the native copy of RNase E will remain, and the other is the last 150 nucleotides of RNase E (from the transformed plasmid) will replace the native copy of RNase E. The result of the desired cross-overs is only the last 150 nucleotides of RNase E will exist in the native locus, and the strain will be resistant to sucrose and sensitive to hygromycin.

M. smegmatis electrocompetent cell preparation

The SS-M_0461 strain of *M. smegmatis* was grown in Difco™ Middlebrooks 7H9 broth with 10% of 10X ADC, shaking at 200 RPM at 37°C to an OD_{600} of 0.8 – 1.0. The cells were then collected by centrifuging at 3900 RPM for 5 minutes at 4°C and discarding the supernatant. The pelleted cells were then washed by gently resuspending in 15 mL of cold 10% glycerol. The cells were collected again by repeating centrifugation at 3900 RPM for 5 minutes at 4°C, discarding supernatant, and washing gently with 5 mL of cold 10% glycerol. A final collection was completed by centrifuging at 3900 RPM for 5 minutes at 4°C, discarding the supernatant, and resuspending gently in 1.5 mL of cold 10% glycerol. The competent cells were split into 3 tubes with 500 µL volume each and stored at -80 °C until use in electroporation.

M. smegmatis transformation

For each transformation, a tube of SS-M_0461 electrocompetent cells were thawed on ice, and 200 ng of plasmid was added to cells. The cells were transferred to a 0.2 cm electroporation cuvette and the cells electroporated with 2.5 kV. The cells were then transferred to a 15 mL conical tube with 2 mL of 7H9 broth supplemented with 0.2% glycerol, 0.05% Tween, and Albumin Dextrose Catalase (ADC with concentration of 5 g/L bovine serum albumin fraction V, 2 g/L 108 dextrose, 0.85 g/L sodium chloride, and 3 mg/L catalase) and incubated at 200 RPM at 37°C for 4 hours. After 4 hours of incubation, the cells were plated on Middlebrooks 7H10 agar plates + 10% ADC and 0.5% glycerol, and grown for 72 hours at 37°C.

Genomic DNA Extraction

For each genomic prep, an overnight culture of *M. smegmatis* was grown at 200 RPM, 37°C to an OD_{600} of 0.8 – 1.0 and centrifuged at 3,900 RPM for 5 minutes. The supernatant was discarded, washed with 200 µL of TE buffer, and transferred to tube pre-filled with 100 µm zirconium beads tube. 400 µL of PCI (Phenol /Chloroform /Isoamyl alcohol; 25:24:1, respectively) was added to the tube and vortexed for 30 seconds. The tube was then centrifuged at 14,000 RPM for 5 minutes, and the upper aqueous layer was collected. 1 mL of ice-cold 95% ethanol was mixed with the aqueous layer and centrifuged at 14,000 RPM for 5 minutes at 4°C. The supernatant was discarded, the pellet was washed with 1 mL of 70% ethanol and centrifuged at 14,000 RPM. After all the supernatant was removed (incubation at room temperature for 5 minutes was occasionally necessary to remove excess supernatant), 100 µL of H_2O was added and the extracted genomic DNA, which was thenstored at room temperature.

M. smegmatis Protein Extraction

An overnight culture of *M. smegmatis* was grown at 200 RPM, 37°C to an OD_{600} of 0.8 – 1.0. The cells were centrifuged at 3,900 RPM for 10 minutes for the 1st wash, and 5 minutes for the 2nd and 3rd wash. All washes were performed using 7H9 broth and the supernatant from each wash was discarded. The cells were then resuspended in 1 mL of lysate buffer (3.5 mL of lysate buffer was made with 1.75 mL of 2X buffer G, 87.5 µL of Triton X-100, 0.7 mL of 50% glycerol, 35 µL of PMSF, 35 µL of 10% SDS, 350 µL of protease buffer, and 542.5 µL of H_2O). The cells were then transferred to a tube and lysed using FastPrep-24 5G Homogenizer from MP Biomedicals for 4 cycles (6.5 m/s, 30 seconds per cycle). The cells were incubated on ice for 1 minute between each cycle. After 4 cycles the cells were centrifuged at 15,000 RPM for 10 minutes at room temperature. The upper aqueous layer was collected and stored at -20°C.

Bicinchoninic Acid (BCA) Assay

BSA standards were set up in a 96 well plate as indicated in Table 4.

Table 4: Preparation of BSA standards.

Vial	Lysate buffer used in protein extraction (µL)	BSA (µL)	Concentration (µg/mL)
A	0	60 (from stock)	2000
B	30	90 (from stock)	1500
C	60	60 (from stock)	1000
D	30	30 (from vial B)	750
E	60	60 (from vial C)	500
F	60	60 (from vial E)	250
G	60	60 (from vial F)	125
H	48	12 (from vial G)	25
I	60	0	0 (blank)

Each standard and protein sample was assayed in duplicate with 25 µL in each well. Each replicate was mixed with 200 µL of BCA reagent (ThermoFisher Scientific). After the reagent was added, the plate was incubated at 37°C for 30 minutes, and then absorbances at 544 nm were collected.

SDS-PAGE

A Mini-PROTEAN TGX Stain-Free Gel from BIO-RAD was assembled in the gel cassette and filled with running buffer (30.3 g of Tris-Base, 144.1 g of glycine, and 10 g of SDS mixed with 1L of H_2O) to make sure the gel cassette did not leak. 2 mg of each sample was prepared, mixed with 4X loading dye in a total volume of 32 µL, and incubated at 65°C for 20 minutes. 10 µL of protein ladder and 32 µL of incubated samples were loaded and ran 140V until the ladder reached the bottom end of the gel.

Western Blotting

The gel from SDS-PAGE was taken out and agitated in 1X transfer buffer (90 mL of H_2O mixed

with 10 mL of 10X transfer buffer made by 3.63 g of Tris-Base and 17.28 g of glycine mixed with 240 mL of 100% methanol and 950 mL) for 20 minutes. A sheet of polyvinylidene difluoride (PVDF) membrane was soaked in 100% methanol for 10 minutes and washed with transfer buffer for 5 minutes. In addition, two filter paper sheets and two foam pads were soaked in transfer buffer for 20 minutes.

The gel was assembled in the blotting cassette in the following order- Cassette Top (Black)//Foam Pad//Filter Paper//Gel//Membrane//Filter Paper//Foam Pad//Cassette Bottom- and was run at 100V for one hour. The membrane was then removed and checked to confirm the ladder was successfully blotted to the membrane. The membrane was then agitated in H_2O for 5 minutes and then blocking agent (100 mL 1X Tris-buffered saline (TBS) with 3 g of non-fat milk) for one hour. The membrane was next washed in 1X TBS twice, for 5 minutes, and agitated overnight with primary antibody (1:2000 ratio of Monoclonal ANTI-FLAG M2 antibody in blocking agent) at 4°C.

The primary antibody was then removed, washed with 1X TBS for 5 minutes, and agitated in secondary antibody (30 mL of blocking agent with 1 µL of Goat Anti-Mouse HRP antibody) for 30 minutes at room temperature. The membrane was washed with 1X TBS with 0.05% Tween 20 for five times, 5 minutes each. 2 mL of enhanced chemiluminescence (ECL) substrate was spread on the membrane, the membrane was wrapped in plastic wrap, covered to make sure the membrane was not exposed to light, and brought to a dark room. The X-ray film was exposed for 2 seconds with the membrane and imaged by Azure™ cSeries Advanced Imaging Systems.

Coomassie Blue Staining

The gel from SDS-PAGE was removed and agitated in H_2O for 5 minutes. Next, the gel was agitated in staining solution (50% methanol, 10% acetic acid, and 0.25% Coomassie brilliant blue) for 30 minutes. The gel was then washed three times with destain solution (50% methanol and 10% acetic acid) for 30 minutes, and was then agitated overnight in low methanol destain solution (5% methanol and 7.5% acetic acid). The stained gel was imaged using Gel Doc™ XR+ Gel Documentation System.

RESULTS

In order to map the *M. smegmatis* RNase E domain structure and identify potential scaffold domains, RNase E amino acid sequences from several species related to *M. smegmatis* were aligned with *M. smegmatis* RNAse E sequence as the template.

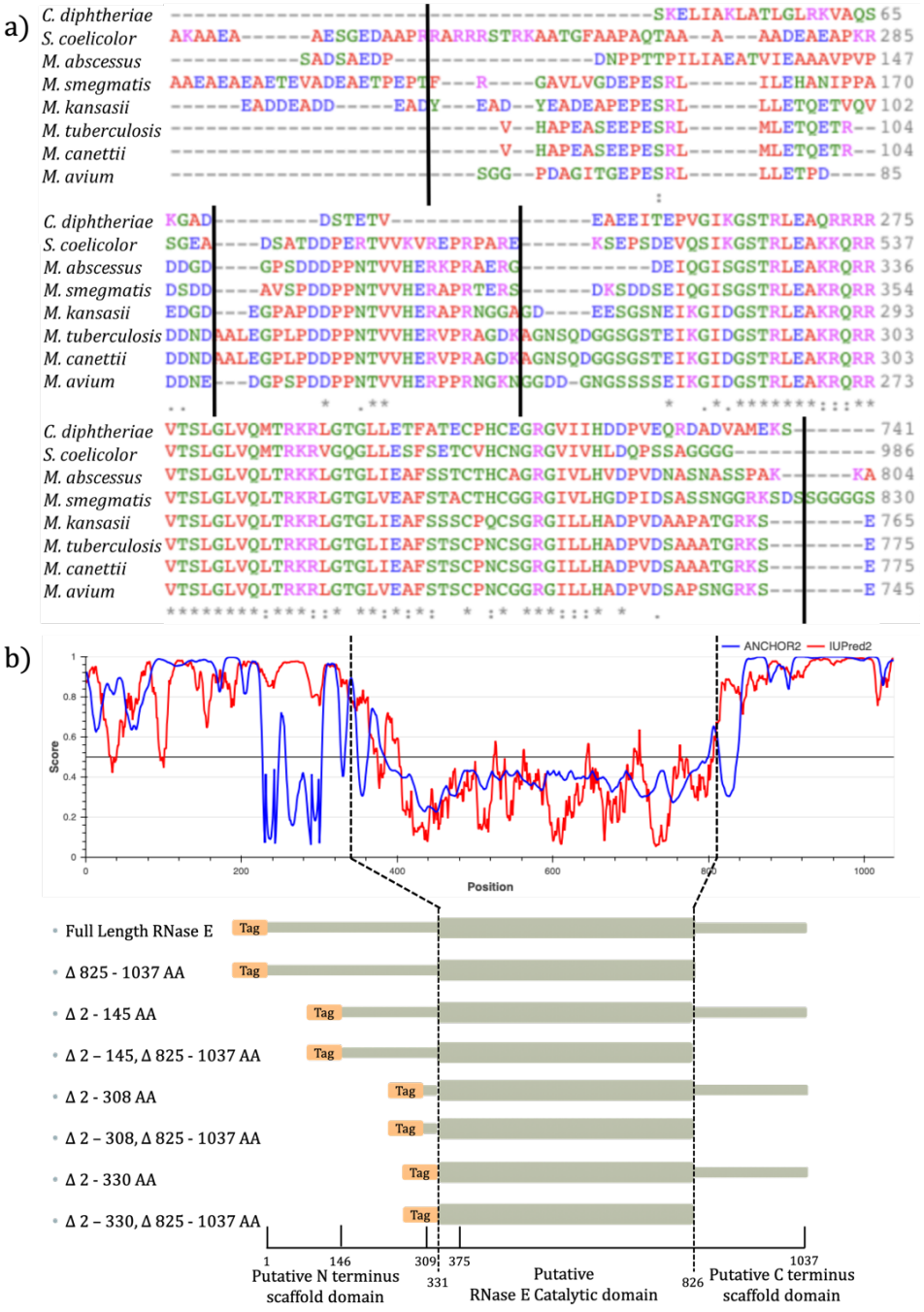

Figure 1. Choosing deletion boundaries in *M. smegmatis rne.* a) Sequence alignment using *M. smegmatis* RNase E with closely-related species (Retrieved from https://www.ebi.ac.uk/Tools/msa/clustalw2/). Each line represents the deletion boundary. b) Disorder and binding prediction (based upon amino acid sequence) for *M. smegmatis* RNase E with schematic of deletions. (IUPred2A: https://iupred2a.elte.hu/) Red predicts disorder, where values over 0.5 are predicted to be disordered regions. Blue predicts the possible binding proteins binding regions,; values over 0.5 were predicted to be able to act as binding domains.

The alignment shows low similarity exists in the first 145 and beyond 825 amino acids (AAs). It can be predicted these two regions might represent binding domains for species-specific partners. Amino acid sequence from 331-824 has high similarity, and thus is expected to represent the catalytic region. Moderate similarity between 309 and 330 AAs is observed, and thus could represent sequence within either the binding or catalytic domains. Deletion boundaries were chosen to account for each of these putative domains (i.e., binding, catalytic, and undetermined) (See Fig 1 a).

To clarify and support the definition of these deletion regions, disorder and binding prediction for *M. smegmatis* RNase E was performed (See Fig 1b). Typically catalytic domains are highly ordered, while binding domains show a lower degree of order. Based on the output, most of the hypothesized catalytic domain is predicted to be structured (indicated by a score less than 0.5 in red) rather than in binding other proteins (indicated by a score less than 0.5 in blue). All the truncated *rne* plasmids were constructed using PCR and Gibson assembly.

All the constructed plasmids were confirmed via PCR and sequenced. In order to construct strains which only contained the truncated *rne* in the genome, the integration of truncated *rne* plasmid and removal of the native copy *rne* was necessary. Modification of *M. smegmatis* via integration at the L5 site in the chromosome ris reported to a higher efficiency when compared to integration at Giles and Tweety (Lee et al, 1991). As a result, it was decided to construct a strain of 6X-His and 3X-Flag tagged RNase E (*t-rne*) integrated at the L5 chromosomal site.

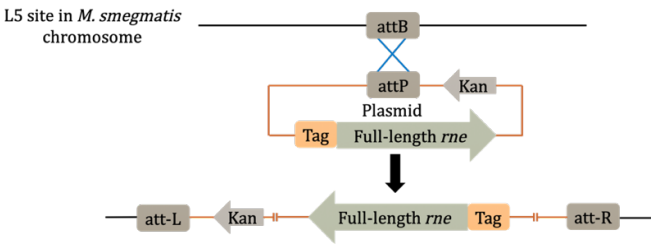

Figure 2. Integration of tagged full-length rne into L5 site in *M. smegmatis*. The *attP* core sequence within the plasmid is a highly similar sequence comparing to the *attB* core sequence at L5 site in *M. smegmatis*. The similarity drives the cross-over event and thereby integrates the plasmid containing tagged full-length RNase E.

To generate the desired strain, a plasmid harboring *t-rne*, the kanamycin resistance gene, and *attP* core sequence was constructed and transformed via electroporation into wild type *M. smegmatis*. The crossover between the plasmid and the L5 site (see Figure 2) is driven by the similarity between the *attP* core sequence in the plasmid and the *attB* core sequence in the L5 site (Lee et al, 1991). All the selected colonies that were confirmed via PCR to have properly integrated *t-rne* at L5 were sequenced.

The next step was to remove the native copy of *rne* from the above strain with *t-rne* integrated at the L5 site, using the 2-Step strategy. A plasmid (pSS299) with only last 150 nucleotides (nts) of *rne* flanked by 1000nt upstream and 850nt downstream of the native *rne* sequence was transformed. The plasmid also contains the hygromycin (Hyg) resistance gene and Sac B which makes sucrose toxic to the cells. Because pSS 299 has the exact same upstream and downstream regions of native copy *rne*, two cross-over possibilities can occur. In order to integrate the plasmid and replace the native *rne* gene, one cross-over must occur upstream and the other downstream of the 5' truncated *rne* sequence. The first and second cross-overs must not happen in the same upstream or downstream region, which essentially reverses the integration event; native *rne* thus remains in the chromosome. (see Figure 3b). Identifying colonies that successfully integrated the truncated, nonfunctional *rne* and removed the native copy was done by screening for *rne* length and sequence by colony PCR followed by DNA sequencing, and Hyg resistance and sucrose sensitivity by plating. The *M. smegmatis* that was confirmed to have *t-rne* at the L5 site and not have the native copy of *rne* was designated asSS-M_0461.

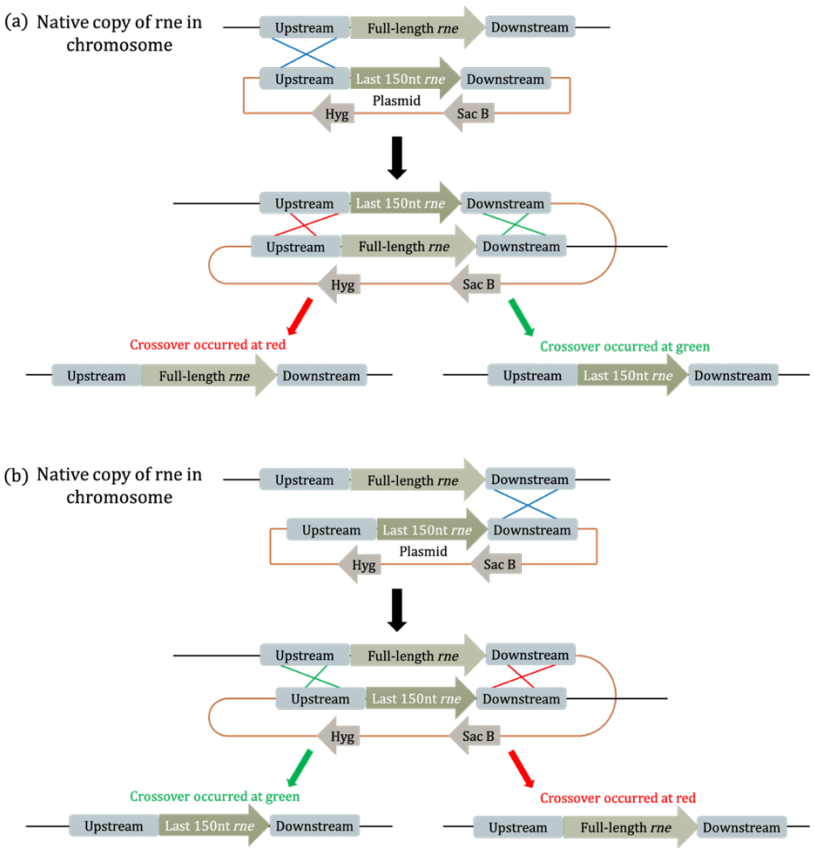

Figure 3. Removing the native copy of rne using 2-step strategy. By transforming plasmid pSS299 which has only the last 150 nucleotides of native *rne* and the exact same upstream and downstream sequences, cross-over will occur twice and have two different outcomes based on whether the first cross-over occurs within the upstream (a) or downstream (b) regions flanking the 150nt *rne* sequence. The desired result is the last 150nt of *rne* successfully replaces the native copy of *rne*, thus effectively knocking out the native *rne* gene.

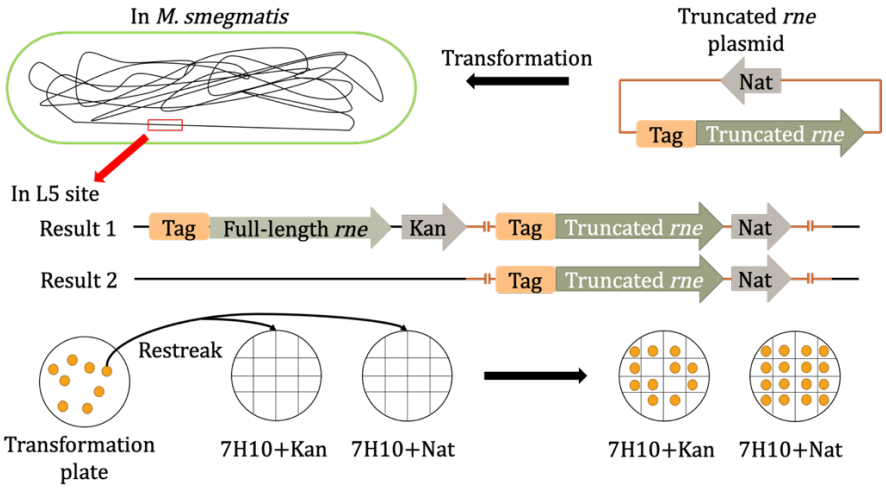

Figure 4. Selection process of modified *rne* in *M. smegmatis* after transformation. Upon transformation with the tagged deletion versions of *rne*, two arrangements exist; (1) the truncated *rne* will replace the full-length version; (2) both full-length and truncated *rne* will present in the genome. To identify those only have the truncated *rne* present in the genome, colonies were re-streaked on 7H10 agar plates with either kanamycin or nourseothricin (Nat) . The colonies that only grew on Nat where chosen for further study.

Base on the sequence alignments and disorder and binding predictions in Figure 1, all truncated *rne* plasmid constructs were made. Upon transformation into SS-M_0461, colonies with modified *rne* in the L5 site were identified by restreaking from the initial transformation plate onto 7H10 agar plates with either kanamycin or nourseothricin (see Figure 4.). If the cells have only the modified *rne* gene in the L5 site, they should grow on plates containing nourseothricin, but will not survive on plates containing kanamycin. Colonies meeting these criteria were checked with PCR and sequencing to confirm the truncated versions were indeed properly integrated.

In addition, while making the truncated *rne* strains, some appeared to yield cultures with highly aggregated cells. To determine if the aggregation phenotype is dominant, each of the truncated *rne* plasmids were transformed into wild-type *M. smegmatis* (See Table 5). Deletion of either of the first 145 or final 213 amino acids either alone or in combination did not yield clumpy strains, and thus were omitted from this analysis.

Table 5: Cell aggregation check.

Strain	Full length RNase E + Truncated RNase E	Truncated RNase E
Full Length	N.A.	Not Clumpy
Δ2-145	N.A.	Not Clumpy
Δ2-145, Δ825-1037	N.A.	Not Clumpy
Δ825-1037	N.A.	Not Clumpy
Δ2-308	Clumpy	Clumpy
Δ2-308, Δ825-1037	Clumpy	Clumpy
Δ2-329	Clumpy	Clumpy
Δ2-329, Δ825-1037	Clumpy	Clumpy

However, cultures grown from strains in which harbored deletions of first 308 ors, 329 AAs alone or in combination with the last 213 AAs appeared to be clumpy. In addition, all yield clumpy cultures when both full-length *rne* and truncated *rne* present are in the genome; thus the clumpy phenotype appearsto be dominant. The region that might be responsible for cell aggregation might be between 146-308 AAs, as the strain with the first 145 AAs deleted did not appear clumpy.

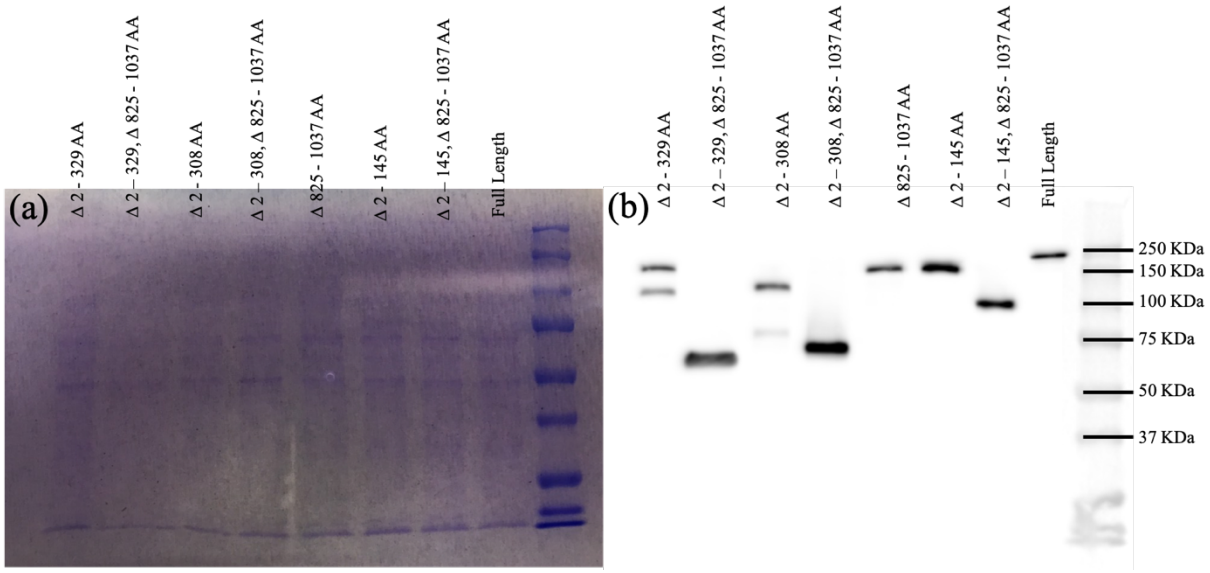

Figure 5. SDS_PAGE and Western blotting for RNase E variants. (a) Coomassie Blue staining showing protein abundance for all samples; 2 mg of total protein was added to each lane. (b) Western blotting using Anti-FLAG primary antibody detecting RNase E variants.

Proteins were extracted from each strain cultures and sequence-confirmed truncated *rne* strain, and SDS-PAGE was performed. The sizes of each truncated RNase E version were compared upon Western blotting (Fig. 5). By taking the full length as a comparison, it can be seen that the first 308 AAs deletion and first 329 AAs deletion, other truncated *RNase E* had the correct relative sizes which implied that the truncated *RNase E* strains can express properly.

For first 308 AAs deletion sample, it appears some degradation occurs as there is a faint band located at a position at about 75 KDa. For first 329 AAs deletion sample, though the lower band seems to have a similar size with the first 308 AAs deletion, but there is a band that also has similar size with the full length RNase E, which implies this strain might contain both versions in the genome.

CONCLUSION & DISCUSSION

To map the *M. smegmatis* RNase E domain structure and identify potential scaffold domains, RNase E sequence alignment and disorder prediction results were obtained. The putative region for catalytic domain of RNase E in *M. smegmatis* is 493 amino acids long, from 331-824. In addition, when the putative N-terminal scaffold domain of *RNase E* was truncated beyond the first 308 amino acids, the cells appeared to be more aggregated; this aggregation was found to be dominant as the strain appeared to be clumpy when both the native and truncated copies of *rne* were present in the genome.

Other than the first 329 AAs deletion (Δ 2 - 329 AA) strain, all other truncated *rne* strains were successfully made. In the western blot the strain with the first 308 AAs deleted (Δ 2 - 308 AA) seems to have some degradation, though the correct size is also observed in the blot. For the first 329 AAs deletion strain (Δ 2 - 329 AA), it appears both full length and truncated tagged *rne* exists in the genome, as it has a band that looks the same as the full length and another band looks similar to the first 308 AAs deletion (Δ 2 - 308 AA).

As a result, it is suggested to double-check the sequence and if necessary, repeat the electroporation and integration steps. After having at least one candidate for all the truncated RNase E strains, the future experiments will be identifying protein candidates using mass spectrometry, and based on the candidate results use biochemical approaches to test putative protein-protein interactions. Another experiment is to test that whether truncated RNase E will affect mRNA half-life in *M. smegmatis*.

REFERENCES

1. World Health Organization (WHO) (2018, October 05). Global tuberculosis report 2018. Retrieved from http://www.who.int/tb/publications/global_report/en/
2. Centers for Disease Control and Prevention (CDC) (2018, August 17). Tuberculosis (TB). Retrieved from https://www.cdc.gov/tb/topic/treatment/tbdisease.html
3. Belasco, J. G., & Biggins, C. F. (2003, January 16). Mechanisms of mRNA decay in bacteria: A perspective. Retrieved from https://www.sciencedirect.com/science/article/pii/0378111988901230
4. Kovacs, L., Csanadi, A., Megyeri, K., Kaberdin, V. R., & Miczak, A. (2013, November 14). Mycobacterial RNase E-Associated Proteins. Retrieved from https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1348-0421.2005.tb03697.x
5. Mackie, G. A. (2013, January). RNase E: At the interface of bacterial RNA processing and decay. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/23241849
6. Hammarlöf, D. L., Bergman, J. M., Garmendia, E., & Hughes, D. (2015, October). Turnover of mRNAs is one of the essential functions of RNase E. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/26094815
7. Vanzo NF, Li YS, Py B, Blum E, Higgins CF, (1998). Ribonuclease E organizes the protein interactions in the Escherichia coli RNA degradosome. Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC317140/
8. Coburn, G. A., Miao, X., Briant, D. J., & Mackie, G. A. (1999, October 01). Reconstitution of a minimal RNA degradosome demonstrates functional coordination between a 3' exonuclease and a DEAD-box RNA helicase. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/10521403
9. Mohanty, B. and Kushner, S. (2000). Polynucleotide phosphorylase functions both as a 3' right-arrow 5' exonuclease and a poly(A) polymerase in Escherichiacoli. [online] Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC17278/
10. Anderson, K. L., & Dunman, P. M. (2009, March). Messenger RNA Turnover Processes in Escherichia coli, Bacillus subtilis, and Emerging Studies in Staphylococcus aureus. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/19936110
11. Kime, L., Jourdan, S. S., Stead, J. A., Hidalgo-Sastre, A., & McDowall, K. J. (2010, May). Rapid cleavage of RNA by RNase E in the absence of 5´ monophosphate stimulation. Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2948425/
12. Miczak, A., Kaberdin, V. R., Wei, C. L., & Lin-Chao, S. (1996, April 30). Proteins associated with RNase E in a multicomponent ribonucleolytic complex. Retrieved from

https://www.ncbi.nlm.nih.gov/pmc/articles/PMC39450/

13. Murashko, O. N., Kaberdin, V. R., & Lin-Chao, S. (2012, May 01). Membrane binding of Escherichia coli RNase E catalytic domain stabilizes protein structure and increases RNA substrate affinity. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/22509045
14. Voss, J. E., Luisi, B. F., & Hardwick, S. W. (2014, December 01). Molecular recognition of RhlB and RNase D in the Caulobacter crescentus RNA degradosome. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/25389270
15. Khemici, V., Poljak, L., Luisi, B. F., & Carpousis, A. J. (2008, November). The RNase E of Escherichia coli is a membrane-binding protein. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/18976283
16. Lehnik-Habrink, M., Pförtner, H., Rempeters, L., Pietack, N., Herzberg, C., & Stülke, J. (2010, August). The RNA degradosome in Bacillus subtilis: Identification of CshA as the major RNA helicase in the multiprotein complex. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/20572937
17. Płociński, P., Macios, M., Houghton, J., Niemiec, E., Płocińska, R., Brzostek, A., . . . Dziembowski, A. (2019, April 08). Proteomic and transcriptomic experiments reveal an essential role of RNA degradosome complexes in shaping the transcriptome of Mycobacterium tuberculosis. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/30957850
18. Kovacs, L., Csanadi, A., Megyeri, K., Kaberdin, V. R., & Miczak, A. (2005). Mycobacterial RNase E-associated proteins. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/16301812
19. Csanadi, A., Faludi, I., & Miczak, A. (2009, November). MSMEG_4626 ribonuclease from Mycobacterium smegmatis. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/19153821